NAAN MUDHALVAN

DATA SCIENCE

PROJECT TITLE

Covid-19 Vaccines Analysis

REG NO:712221104005

NAME:MEGHA.M.R

DEPT:COMPUTER SCIENCE AND ENGINEERING YEAR&SEM:III&05

COLLEGE:PARK COLLEGE OF ENGINEERING AND TECHNOLOGY


DATA
SCIENCE

Problem Definition :

- **The problem is to conduct an in-depth analysis of Covid-19 vaccine data, focusing on vaccine efficacy, distribution, and adverse effects. The goal is to provide insights that aid policymakers and health organizations in optimizing vaccine deployment strategies. This project involves data collection, data preprocessing, exploratory data analysis, statistical analysis, and visualization.**

Design Thinking

- Data Collection: Collect customer data, including attributes like purchase history, demographic information, and interaction behavior.
- Data Preprocessing: Clean and preprocess the data, handle missin values, and convert categorical features into numerical representations.
- DesignClustering Algorithms: Apply clustering algorithms like K-Means,

Data set:

https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress